On the representation of Fleming-Viot models from a Bayesian perspective

MATTEO RUGGIERO
University of Pavia, Italy

Joint with STEPHEN G. WALKER, University of Kent, UK

YEP VI, 2009, Eurandom

Outline

Dirichlet process

Definition. (Ferguson, 1973) Given a finite non-null measure α on $(\mathcal{X}, \mathscr{X})$, a random probability measure μ on $(\mathcal{X}, \mathscr{X})$ is said to be a Dirichlet process with parameter α, denoted $\mu \sim \mathscr{D}_\alpha$, if for every $K = 1, 2, \dots$ and for every measurable partition $(B_1, \dots, B_K)$ of $\mathscr{X}$,

$$\big(\mu(B_1), \dots, \mu(B_K)\big) \sim \mathrm{Dirichlet}\big(\alpha(B_1), \dots, \alpha(B_K)\big).$$

Conjugate posterior process. (Ferguson, 1973) Let μ be a Dirichlet process on $(\mathcal{X}, \mathscr{X})$ with parameter α, and let $X_1, \dots, X_n$ be a sample of size n from μ. Then

$$\mu \mid X_1, \dots, X_n \sim \mathscr{D}_{\alpha + \sum_{i=1}^n \delta_{X_i}}$$

where δ_x denotes a point mass at x.

Connection with GEM distribution. Sethuraman (1994) proposed a series representation for the Dirichlet process, the so-called *stick-breaking construction*, with locations $Y_i \overset{iid}{\sim} \alpha/\alpha(\mathcal{X})$ and weights p_i with GEM distribution

$$\mathscr{D}_\alpha = \mathcal{L}\Bigg(\sum_{i=1}^{\infty} \underbrace{\Bigg(V_i \prod_{j=1}^{i-1}(1 - V_j)\Bigg)}_{p_i} \delta_{Y_i}\Bigg) \qquad V_i \overset{iid}{\sim} \mathrm{Beta}(1, \alpha(\mathcal{X})),\ Y_i \overset{iid}{\sim} \frac{\alpha}{\alpha(\mathcal{X})}$$

A Polya urn for the Dirichlet process

Blackwell and MacQueen (1973)

Let again α be a finite measure on $(\mathcal{X}, \mathscr{X})$, and let $\{X_n\}_{n\geq 1}$ be such that

$$X_1 \sim \frac{\alpha}{\alpha(\mathcal{X})}, \qquad X_{n+1}|X_1,\dots,X_n \sim \frac{\alpha + \sum_{i=1}^n \delta_{X_i}}{\alpha(\mathcal{X})+n}, \qquad n > 1.$$

The observed colours have higher probability of being drawn again. If α is non atomic, there is a continuum of colours. Then $\{X_n\}_{n\geq 1}$ is called Polya sequence with parameter α and

(a) $\dfrac{\alpha + \sum_{i=1}^n \delta_{X_i}}{\alpha(\mathcal{X})+n} \Longrightarrow \mu^* \;\; a.s.,$ $\qquad \mu^*$ being a discrete measure

(b) $\mu^* \sim \mathscr{D}_\alpha$

(c) $X_1, X_2, \dots \mid \mu^* \overset{iid}{\sim} \mu^*$

Rephrasing, we can write the joint law of $X_1,\dots,X_n$, for $n \geq 1$, as

$$\mathbb{P}(X_1 \in \mathrm{d}x_1,\dots,X_n \in \mathrm{d}x_n) = \mathbb{E}\Bigg[\prod_{i=1}^n \mu(\mathrm{d}x_i)\Bigg] = \int_{\mathscr{F}} \prod_{i=1}^n \mu(\mathrm{d}x_i)\mathscr{D}_\alpha(\mathrm{d}\mu)$$

where $\mathscr{D}_\alpha$ is called de Finetti measure of the sequence $X_1, X_2, \dots$

Gibbs sampling

Geman and Geman (1984)

Special case of a Metropolis-Hastings algorithm, broadly used in Bayesian inference.

Suppose we want to sample from some joint distribution $p_{X,Y}(x,y)$, but this is unfeasible. Given the initial value (x_0, y_0), it is usually easier to sample from the *(full) conditional distributions*

$$\begin{aligned} X_1 &\sim p_{X|Y}(x \,|\, y_0) \\ Y_1 &\sim p_{Y|X}(y \,|\, x_1) \\ X_2 &\sim p_{X|Y}(x \,|\, y_1) \\ &\vdots \end{aligned}$$

and so on. Then $\{(x_n, y_n)\}_{n\geq 1}$ is a Markov chain with stationary distribution $p_{X,Y}(x,y)$.

Taking M such chains $\{(x_n^i, y_n^i)\}_{n\geq 1, i=1,\dots,M}$, for sufficiently large $N \geq 1$ we can approximate

$$\frac{1}{M}\sum_{i=1}^{M} f(x_N^i, y_N^i) \approx \int f(x,y) p_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$

If the coordinates are updated in a random order and visited infinitely often, the chain is also reversible w.r.t. $p_{X,Y}(x,y)$.

Generalisation to d-variate case is straightforward.

Outline

Fleming-Viot processes

Fleming and Viot (1979)

A Fleming-Viot process is a probability-measure-valued diffusions which describes the evolution in time of an infinite population subject to *mutation*, *resampling* and (possibly) *selection* and *recombination*.

Among its main features:

- individuals are labeled by points in a complete separable metric space $\mathcal{X}$, called *type space* (for simplicity we assume $\mathcal{X}$ is compact);
- it takes values on the set $\mathscr{P}(\mathcal{X})$ of Borel probability measures;
- it has sample-paths in the space $C_{\mathscr{P}(\mathcal{X})}([0,\infty))$ of continuous functions from $[0,\infty)$ to $\mathscr{P}(\mathcal{X})$.

The neutral version has infinitesimal generator

$$\mathcal{A}_0\varphi(\mu) = \sum_{i=1}^{m}\langle P_i f, \mu^m\rangle + \frac{1}{2}\sum_{1\le k\neq i\le m}\langle \Phi_{ki}f - f, \mu^m\rangle \qquad \langle f,\mu\rangle = \int f \mathrm{d}\mu$$

with domain

$$\mathscr{D}(\mathcal{A}_0) = \left\{\varphi(\mu)\in B(\mathscr{P}(\mathcal{X})) :\ \varphi(\mu) = \langle f,\mu^m\rangle,\ f\in C(\mathcal{X}^m),\ m\in\mathbb{N}\right\}$$

where P is the generator of a Feller mutation process on $\mathcal{X}$, P_i acts on x_i in $f(x_1,\ldots,x_n)$, Φ_{ki} changes x_k to x_i in f.

When

$$Pf(x) = \frac{\theta}{2}\int \big[f(y) - f(x)\big]\nu_0(\mathrm{d}y) \qquad\qquad \nu_0 \text{ non atomic}, \;\; \theta > 0$$

its stationary distribution is $\mathscr{D}_\alpha$, with $\alpha = \theta\nu_0$. (Ethier and Kurtz, 1986)

Its transition function is given by

$$P(t,\mu,\mathrm{d}\nu) = \sum_{m=0}^{\infty} d_m(t)\int_{\mathcal{X}^m} \mathscr{D}_{\alpha+\sum_{i=1}^m \delta_{X_i}}(\mathrm{d}\nu)\mu(\mathrm{d}X_1)\ldots\mu(\mathrm{d}X_m)$$

where $d_m(t) = P(D_t = m)$ and D_t is a death process starting a.s. from ∞, and $\mathscr{D}_{\alpha+\sum_{i=1}^m \delta_{X_i}}$ is a posterior Dirichlet process. (Ethier and Griffiths, 1993)

If we add selection, then the FVP has generator

$$\mathcal{A}_\sigma\varphi(\mu) = \sum_{i=1}^m \langle P_i f, \mu^m\rangle + \frac{1}{2}\sum_{1\le k\ne i\le m}\langle \Phi_{ki} f - f, \mu^m\rangle + \sum_{i=1}^m \langle \sigma_i(\cdot) f - \sigma_{m+1}(\cdot) f, \mu^{m+1}\rangle$$

where $\sigma_i(\cdot) = \sigma(x_i)$ is the selection coefficient, and stationary distribution proportional to

$$e^{2\langle\sigma,\mu\rangle}\,\mathscr{D}_\alpha(\mathrm{d}\mu)$$

where $\langle\sigma,\mu\rangle = \int \sigma(x)\mu(\mathrm{d}x)$. (Ethier and Kurtz, 1994)

Outline

Gibbs sampling the Polya urn

Given an exchangeable vector $\mathbf{X}_n = (X_1, \ldots, X_n)$, define a Gibbs sampler driven Markov chain $\{\mathbf{X}_n(k)\}_{k\geq 1}$ such that at each transition

- x_i is removed from $\mathbf{x}_n = (x_1, \ldots, x_n)$ with probability $1/n$
- a replacement X_i' is sampled from the Blackwell-MacQueen prediction scheme with $\alpha = \theta\nu_0$, where $\theta = \alpha(\mathcal{X})$ and $\nu_0 = \alpha/\alpha(\mathcal{X})$ non atomic, namely

$$X_i' \,|\, \mathbf{x}_{(-i)} \sim \frac{\theta}{\theta+n-1}\nu_0(\mathrm{d}x_i') + \frac{1}{\theta+n-1}\sum_{k\neq i}^{n} \delta_{x_k}(\mathrm{d}x_i') \tag{1}$$

- the arrival state is $(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$.

This amounts to performing a Gibbs sampler on $(X_1, \ldots, X_n)$, with a *random scan* (update X_i with index i random) and *full conditionals* $\mathcal{P}^{\alpha}_{X_i|X_{(-i)}}(\mathrm{d}x_i|x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ given by (1).

This produces a reversible Markov chain $\{(X_1(k), \ldots, X_n(k))\}_{k\geq 1}$ with stationary distribution

$$\mathcal{P}^{\alpha}_{X_1,\ldots,X_n}(\mathrm{d}x_1, \ldots, \mathrm{d}x_n) = \nu_0(\mathrm{d}x_1)\, \frac{\theta\nu_0(\mathrm{d}x_2) + \delta_{x_1}}{\theta+1} \cdots \frac{\theta\nu_0(\mathrm{d}x_n) + \sum_{k=1}^{n}\delta_{x_k}(\mathrm{d}x_n)}{\theta+n-1}$$

The particle process

Embed it in continuous time in $D_{\mathcal{X}^n}[0,\infty)$, with $\mathrm{Exp}(\lambda_n)$ sojourn times, and let

$$\lambda_n = n(\theta + n - 1)/2$$

Remarks

a) λ_n substitutes time rescaling.
b) for $\theta = 0$ there is no mutation, and λ_n is the transition rate of Kingman's coalescent.

The generator of the $\mathcal{X}^n$-valued process is

$$A^n f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\lambda_n}{n(\theta + n - 1)} \int \Big[f(\eta_i(\mathbf{x}|y)) - f(\mathbf{x}) \Big] (\theta\nu_0 + \sum_{k\neq i}^{n} \delta_{x_k})(\mathrm{d}y)$$

where $\eta_i(\mathbf{x}|y) = (x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$.

Define the process of empirical measures $\{\mu_n(t)\}_{t\geq 0} := \{\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i(t)}\}_{t\geq 0}$ with *càdlàg* sample-paths in $D_{\mathscr{P}(\mathcal{X})}[0,\infty)$.

Convergence and stationarity

Neutral diffusion model

Define

$$\varphi_m(\mu) = \langle f, \mu^{(m)} \rangle, \qquad \mu^{(m)} = \frac{(n-m)!}{n!} \sum_{1 \le i_1 \ne \ldots \ne i_m \le n} \delta_{(x_{i_1}, \ldots, x_{i_m})}$$

Then $\mathcal{A}^n \varphi_m(\mu) = \langle A^n f, \mu^{(m)} \rangle$ is the generator of the measure-valued process and

$$||\mathcal{A}^n \varphi_m(\mu) - \mathcal{A}_0 \phi_m(\mu)|| \underset{n \to \infty}{\longrightarrow} 0 \qquad \phi_m(\mu) = \langle f, \mu^m \rangle$$

where $\mathbb{A}$ is the generator of a FV process. Since the linear span of functions ϕ_m is a core for $\mathbb{A}$ in $C(\mathscr{P}(\mathcal{X}))$, and both $\mathbb{A}^n$ and $\mathbb{A}$ generate strongly continuous contraction semigroups, this implies $\{\mu_n(t)\} \Longrightarrow \{\mu_\infty(t)\}$ in $D_{\mathscr{P}(\mathcal{X})}[0, \infty)$, where $\{\mu_\infty(t)\}$ is a FV process.

From de Finetti's theorem, w.p. 1 we have $\mu_n(t) \Rightarrow \mu_\infty(t)$ for every t, and $\mu_\infty(t) \sim \mathscr{D}_{\theta\nu_0}$. From the well-posedness of the martingale problem for $\mathcal{A}_0$, it follows that the stationary distribution of $\{\mu_\infty(t)\}$ is a Dirichlet process $\mathscr{D}_{\theta\nu_0}$.

Outline

A generalised Polya urn scheme

Consider the exchangeable law

$$\mathcal{Q}^{\alpha,\beta_n}_{X_1,\dots,X_n}(\mathrm{d}x_1,\dots,\mathrm{d}x_n) \propto \mathcal{P}^{\alpha}_{X_1,\dots,X_n}(\mathrm{d}x_1,\dots,\mathrm{d}x_n)\beta_n(x_1)\dots\beta_n(x_n) \qquad (2)$$

where we assume $\beta_n \in B(\mathcal{X})$ for all n.

Remark

It can be shown that $\mathcal{Q}^{\alpha,\beta_n}_{X_1,\dots,X_n}$ admits representation in terms of a Dirichlet process mixture (Lo, 1984), a model widely used for Bayesian density estimation.

From (2) the predictive law for x_i is

$$\mathcal{Q}^{\alpha,\beta_n}_{X_i|X_{(-i)}}(\mathrm{d}x_i|x_1,\dots,x_{i-1},x_{i+1},\dots,x_n) \;\propto\; \theta\beta_n(x_i)\,\nu_0(\mathrm{d}x_i) + \sum_{l\neq i}^{n}\beta_n(x_l)\,\delta_{x_l}(\mathrm{d}x_i)$$

and it is clear that for $\beta_n(x) \equiv 1$ it reduces to the Blackwell-MacQueen case

$$\mathcal{Q}^{\alpha,1}_{X_1,\dots,X_n}(\mathrm{d}x_1,\dots,\mathrm{d}x_n) \propto \mathcal{P}^{\alpha}_{X_1,\dots,X_n}(\mathrm{d}x_1,\dots,\mathrm{d}x_n) = \prod_{i=1}^{n}\frac{\theta\nu_0(\mathrm{d}x_i)+\sum_{l\leq i}\delta_{x_l}(\mathrm{d}x_i)}{\theta+i-1}.$$

Gibbs sampling again

Similarly to the neutral case, define a Markov chain $\{\mathbf{X}_n(k)\}_{k\geq 1}$ such that at each transition

- x_i is removed from $\mathbf{x}_n = (x_1, \ldots, x_n)$ with probability $1/n$
- a replacement is sampled from the generalized Blackwell-MacQueen predictive

$$\mathcal{Q}^{\alpha,\beta_n}_{X_i|X_{(-i)}}(\mathrm{d}x_i|x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \propto \theta\beta_n(x_i)\,\nu_0(\mathrm{d}x_i) + \sum_{l\neq i}^{n} \beta_n(x_l)\,\delta_{x_l}(\mathrm{d}x_i)$$

This produces a chain reversible with respect to $\mathcal{Q}^{\alpha,\beta}_{X_1,\ldots,X_n}(\mathrm{d}x_1,\ldots,\mathrm{d}x_n)$.

Embed it in continuous time in $D_{\mathcal{X}^n}[0,\infty)$ with $\mathrm{Exp}(\lambda_{n,i})$ sojourn times such that

$$\lambda_{n,i} = \frac{1}{2}\,n\left(\theta\int \beta_n(u)\,\nu_0(\mathrm{d}u) + \sum_{l\neq i}\beta_n(x_l)\right)$$

and note that $\beta_n \equiv 1 \;\Rightarrow\; \lambda_{n,i} = n(\theta+n-1)/2$

Remark

$\lambda_{n,i}$ depends on the starting state only, hence $\{\mathbf{X}_n(t)\}_{t\geq 0}$ is still a Markov process.

Convergence

Fleming-Viot process with selection

When the weights in $\mathcal{Q}^{\alpha,\beta_n}_{X_1,\ldots,X_n}$ have form $\beta_n(x) = 1 + \frac{2}{n}\sigma(x)$, with $\sigma \in B(\mathcal{X})$ the particle process has generator

$$A^n_\sigma f(\mathbf{x}) = \sum_{i=1}^n \frac{1}{2}\theta \int \Big[f(\eta_i(\mathbf{x}|y)) - f(\mathbf{x})\Big]\{1 + \frac{2}{n}\sigma(y)\}\nu_0(\mathrm{d}y)$$
$$+\frac{1}{2}\sum_{1\le k\ne i\le n}\Big[f(\eta_i(\mathbf{x}|x_k)) - f(\mathbf{x})\Big] + \frac{1}{n}\sum_{1\le k\ne i\le n}\sigma(x_k)\Big[f(\eta_i(\mathbf{x}|x_k)) - f(\mathbf{x})\Big]$$

Remark

σ represents the fitness of the offspring, acting as fertility selection.

Since $\langle A^n_\sigma f, \mu^{(m)}\rangle \to \mathcal{A}_\sigma\phi_m(\mu)$ strongly, it can be shown that the process of empirical measures converges in distribution in $D_{\mathscr{P}(\mathcal{X})}[0,\infty)$ to the FV process with fertility selection

$$\left\{\frac{1}{n}\sum_{i=1}^n \delta_{x_i(t)},\ t\ge 0\right\} \underset{n\to\infty}{\Longrightarrow} \{\mu^\sigma_\infty(t),\ t\ge 0\} \qquad \text{in } D_{\mathscr{P}(\mathcal{X})}[0,\infty)$$

where μ^σ_∞ has generator $\mathcal{A}_\sigma$.

Diploid case

For a diploid population, take a bivariate selection function $\beta_n(x, y) \in B_{\text{sym}}(\mathcal{X}^2)$ and consider the law, joint with pairings P_n,

$$\mathcal{Q}^{\alpha,\beta_n}_{X_1,\dots,X_n,P_n}(\mathrm{d}x_1, \dots, \mathrm{d}x_n, P_n) \propto \mathcal{P}^{\alpha}_{X_1,\dots,X_n}(\mathrm{d}x_1, \dots, \mathrm{d}x_n) \prod_k \beta_n(x_k, x_{j_k})$$

With appropriate modifications, the same procedure leads to a generalized urn scheme with conditional law proportional to

$$\theta \sum_{j\neq i}^{n} \beta_n(x_i, x_j)\, \nu_0(\mathrm{d}x_i) + \sum_{k\neq i}^{n} \sum_{j\neq i}^{n} \beta_n(x_i, x_j)\, \delta_{x_k}(\mathrm{d}x_i)$$

and a particle process with Poisson rate

$$\lambda_{n,i} = \frac{1}{2} n \left(\theta \int \sum_{j\neq i}^{n} \beta_n(x_i, x_j)\, \nu_0(\mathrm{d}x_i) + \sum_{k\neq i}^{n} \sum_{j\neq i}^{n} \beta_n(x_k, x_j) \right)$$

whose process of empirical measures converge to a FV process with diploid selection.
When $\beta_n(x) = \int \beta_n(x, y)\mu(y)$ we recover the haploid case. When $\beta_n(x, y) \equiv 1$ we recover

$$n(\theta + n - 1)/2 \qquad \text{and} \qquad \theta\nu_0(\mathrm{d}x_i) + \sum_{k\neq i}^{n} \delta_{x_k}(\mathrm{d}x_i).$$

Stationarity

We exploit the representation of $\mathcal{Q}^{\alpha,\beta_n}_{X_1,\dots,X_n}(\mathrm{d}x_1,\dots,\mathrm{d}x_n)$ in terms of Dirichlet process mixture model. Given

$$z_i|x_i \stackrel{ind}{\sim} K_n(\cdot|x_i) \qquad x_i|\mu \stackrel{iid}{\sim} \mu \qquad \mu \sim \mathscr{D}_{\theta\nu_0}$$

so that

$$\mathcal{L}(X_1,\dots,X_n|z_1,\dots,z_n) \propto K_n(\cdot|x_1)\dots K_n(\cdot|x_n)\mathcal{P}^{\alpha}_{X_1,\dots,X_n}.$$

Assuming $K_n(1|x_i)=\beta_n(x_i)$ we have $\mathcal{Q}^{\alpha,\beta_n}_{X_1,\dots,X_n}$ is the stationary of $(\mathbf{x}|\mathbf{z}_n=\mathbf{1})$.

Consider the Gibbs sampler extended to $(x_1,\dots,x_n,\mu|\mathbf{z}_n=\mathbf{1})$, alternating updates to

$$(x_1,\dots,x_n|\mu,\mathbf{z}_n=\mathbf{1}) \qquad \text{and} \qquad (\mu|x_1,\dots,x_n,\mathbf{z}_n=\mathbf{1}).$$

Hence $(\mu|\mathbf{z}_n=\mathbf{1})$ is a MV chain with stationary $\mathcal{L}(\mu|\mathbf{z}_n=\mathbf{1})$. From Bayes' theorem we have

$$\begin{aligned}\mathcal{L}(\mu|\mathbf{z}_n=\mathbf{1}) &\propto \mathcal{L}(\mathbf{z}_n=\mathbf{1}|\mu)\mathscr{D}_{\theta\nu_0}(\mathrm{d}\mu)\\ &\propto\left[\int\beta_n(y)\mu(\mathrm{d}y)\right]^n\mathscr{D}_{\theta\nu_0}(\mathrm{d}\mu)=\Pi_n(\mathrm{d}\mu)\end{aligned}$$

The limit of Π_n will be the de Finetti measure of the sequence $(x_1, x_2, \dots | \mathbf{z}_\infty = \mathbf{1})$, since

$$(x_1, \dots, x_n | \mu, \mathbf{z}_n = \mathbf{1}) \overset{iid}{\sim} \mu \qquad\qquad \mu \sim \Pi_n$$

from which $(x_1, \dots, x_n | \mathbf{z}_n = \mathbf{1}) \sim \mathcal{Q}^{\alpha, \beta_n}_{X_1, \dots, X_n}$ implies

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \Longrightarrow \mu^* \quad a.s. \qquad\qquad \mu^* \sim \Pi_\infty \quad \text{(if it exists)} \tag{3}$$

When $\mathcal{X}$ is compact, $\mathscr{P}(\mathcal{X})$ (with the topology of weak convergence) is compact, hence $\{\Pi_n\}$ is tight, and Π_∞ is well defined. If $\beta_n(x) = 1 + \frac{2}{n}\sigma(x)$ we have

$$\begin{aligned} \Pi_\infty(\mathrm{d}\mu) &\propto \lim_n \left[1 + \frac{2}{n} \int \sigma(y)\mu(\mathrm{d}y)\right]^n \mathscr{D}_{\theta\nu_0}(\mathrm{d}\mu) \\ &\propto e^{2\int \sigma \mathrm{d}\mu} \mathscr{D}_{\theta\nu_0}(\mathrm{d}\mu). \end{aligned}$$

Since the martingale problem for $\mathcal{A}_\sigma$ is well-posed, (3) is enough to conclude that Π_∞ is the stationary distribution of the FV process with selection.

References

ANTONIAK C.E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist. 2.*

BLACKWELL D. AND MACQUEEN J.B. (1973). Ferguson distributions via Polya urn schemes. *Ann. Statist. 1.*

DAWSON D.A. AND GREVEN A. (1999). Hierarchically interacting Fleming-Viot processes with selection and mutation: multiple space time scale analysis and quasi-equilibria. *Electron. J. Probab. 4.*

DAWSON D.A., GREVEN A. AND VAILLANCOURT J. (1995). Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc., 347.*

DONNELLY P. AND KURTZ T.G. (1996). A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab. 24*.

DONNELLY P. AND KURTZ T.G. (1999). Genealogical processes for Fleming-Viot models with selection and recombination. *Ann. Appl. Probab. 9.*

ETHIER S.N. AND GRIFFITHS R.C. (1993). The transition function of a Fleming-Viot process. *Ann. Probab. 21.*

ETHIER S.N. AND KURTZ T.G. (1986). *Markov processes: characterization and convergence. Wiley.*

ETHIER S.N. AND KURTZ T.G. (1994). Convergence to the Fleming-Viot process in the weak atomic topology. *Stoch. Proc. Appl. 54.*

FERGUSON T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist., 1.*

FLEMING W. H. AND VIOT M. (1979). Some measure-valued processes in population genetics theory. *Indiana University Mathematics J. 28.*

GEMAN, S. AND GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Patt. Anal. Mach. Intelligence, 6.*

LO A. Y. (1984) . On a class of Bayesian nonparametric estimates I: density estimates *Ann. Statist., 12.*

SETHURAMAN J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica, 4.*

VAILLANCOURT J. (1990). Interacting Fleming-Viot processes. *Stoch. Proc. Appl., 36.*